

OIPE

## RAW SEQUENCE LISTING

DATE: 12/05/2001

PATENT APPLICATION: US/09/993,179

TIME: 09:54:30

Input Set : A:\sequence listing.txt

Output Set: N:\CRF3\11212001\I993179.raw

Does Not Comply  
Corrected Diskette Needed*pp 1-3*

3 <110> APPLICANT: McCarthy, Sean A.  
 4 Kuranda, Michael Joseph  
 5 Bulawa, Christine Ellen  
 6 Bossone, Steven  
 8 <120> TITLE OF INVENTION: METHOD FOR IDENTIFYING GENES ENCODING SIGNAL SEQUENCES  
 10 <130> FILE REFERENCE: 09404/032001  
 12 <140> CURRENT APPLICATION NUMBER: US/09/993,179  
 13 <141> CURRENT FILING DATE: 2001-11-06  
 15 <160> NUMBER OF SEQ ID NOS: 15  
 17 <170> SOFTWARE: FastSEQ for Windows Version 3.0

## ERRORED SEQUENCES

52 <210> SEQ ID NO: 2  
 53 <211> LENGTH: 50  
 54 <212> TYPE: PRT  
 55 <213> ORGANISM: Homo sapiens  
 57 <400> SEQUENCE: 2  
 58 Met Lys Gly Thr Cys Val Ile Ala Trp Leu Phe Ser Ser Leu Gly Leu  
 E--> 59 1 5 10 15 20 25 30 35 40 45  
 60 Trp Arg Leu Ala His Pro Glu Ala Gln Gly Thr Thr Gln Cys Gln Arg  
 E--> 61 20 25 30  
 62 Thr Leu Glu Val Asn Ile Val Ser Pro Ser Ser Lys Ala Thr Phe Ser  
 E--> 63 35 40 45  
 64 Pro Ser  
 65 50  
 112 <210> SEQ ID NO: 4  
 113 <211> LENGTH: 125  
 114 <212> TYPE: PRT  
 115 <213> ORGANISM: Homo sapiens  
 117 <400> SEQUENCE: 4  
 118 Met Arg Ser Leu Leu Arg Thr Pro Phe Leu Cys Gly Leu Leu Trp Ala  
 E--> 119 1 5 10 15  
 120 Phe Cys Ala Pro Gly Ala Arg Ala Glu Glu Pro Ala Ala Ser Phe Ser  
 E--> 121 20 25 30  
 122 Gln Pro Gly Ser Met Gly Leu Asp Lys Asn Thr Val His Asp Gln Glu  
 E--> 123 35 40 45  
 124 His Ile Met Glu His Leu Glu Gly Val Ile Asn Lys Pro Glu Ala Glu  
 E--> 125 50 55 60  
 126 Met Ser Pro Gln Glu Leu Gln Leu His Tyr Phe Lys Met His Asp Tyr  
 E--> 127 65 70 75 80  
 128 Asp Gly Asn Asn Leu Leu Asp Gly Leu Glu Leu Ser Thr Ala Ile Thr  
 E--> 129 85 90 95  
 130 His Val His Lys Glu Glu Gly Ser Glu Gln Ala Pro Leu Glu Val Asn  
 E--> 131 100 105 110  
 132 Ile Val Ser Pro Ser Ser Lys Ala Thr Phe Ser Pro Ser

*misaligned  
 amino acid nos.  
 (see item 3  
 on Error Summary  
 sheet)*

*same  
 error*

RAW SEQUENCE LISTING  
PATENT APPLICATION: US/09/993,179

DATE: 12/05/2001  
TIME: 09:54:30

Input Set : A:\sequence listing.txt  
Output Set: N:\CRF3\11212001\I993179.raw

E--> 133      115      120      125  
135 <210> SEQ ID NO: 5  
136 <211> LENGTH: 32  
137 <212> TYPE: PRT  
138 <213> ORGANISM: Mus musculus  
140 <400> SEQUENCE: 5  
141 Met Lys Gly Ala Cys Ile Leu Ala Trp Leu Phe Ser Ser Leu Gly Val  
E--> 142      1      5      10      15  
143 Trp Arg Leu Ala Arg Pro Glu Thr Gln Asp Pro Ala Lys Cys Gln Arg  
E--> 144      20      25      30  
146 <210> SEQ ID NO: 6  
147 <211> LENGTH: 45  
148 <212> TYPE: PRT  
149 <213> ORGANISM: Homo sapiens  
151 <400> SEQUENCE: 6  
152 Met Ser Pro Gln Glu Leu Gln Leu His Tyr Phe Lys Met His Asp Tyr  
E--> 153      1      5      10      15  
154 Asp Gly Asn Asn Leu Leu Asp Gly Leu Glu Leu Ser Thr Ala Ile Thr  
E--> 155      20      25      30  
156 His Val His Lys Glu Glu Gly Ser Glu Gln Ala Pro Leu  
E--> 157      35      40      45  
238 <210> SEQ ID NO: 14  
239 <211> LENGTH: 32  
240 <212> TYPE: PRT  
241 <213> ORGANISM: Homo sapiens  
243 <400> SEQUENCE: 14  
244 Met Lys Gly Thr Cys Val Ile Ala Trp Leu Phe Ser Ser Leu Gly Leu  
E--> 245      1      5      10      15  
246 Trp Arg Leu Ala His Pro Glu Ala Gln Gly Thr Thr Gln Cys Gln Arg  
E--> 247      20      25      30  
249 <210> SEQ ID NO: 15  
250 <211> LENGTH: 108  
251 <212> TYPE: PRT  
252 <213> ORGANISM: Homo sapiens  
254 <400> SEQUENCE: 15  
255 Met Arg Ser Leu Leu Arg Thr Pro Phe Leu Cys Gly Leu Leu Trp Ala  
E--> 256      1      5      10      15  
257 Phe Cys Ala Pro Gly Ala Arg Ala Glu Glu Pro Ala Ala Ser Phe Ser  
E--> 258      20      25      30  
259 Gln Pro Gly Ser Met Gly Leu Asp Lys Asn Thr Val His Asp Gln Glu  
E--> 260      35      40      45  
261 His Ile Met Glu His Leu Glu Gly Val Ile Asn Lys Glu Ala Glu Met  
E--> 262      50      55      60  
263 Ser Pro Gln Glu Leu Gln Leu His Tyr Phe Lys Met His Asp Tyr Asp  
E--> 264      65      70      75      80  
265 Gly Asn Asn Leu Leu Asp Gly Leu Glu Leu Ser Thr Ala Ile Thr His  
E--> 266      85      90      95  
267 Val His Lys Glu Glu Gly Ser Glu Gln Ala Pro Leu  
E--> 268      100      105

*same*

*same*

*same*

*same*

*(see next page)*

09/993/79

3

<400> 1  
 ggggaccgtg tttgtggccc ccaagccggt gccccccatt ttggaactca gcgagtaggg 60  
 ggcggctctg gggaagtggc agggggcgca gcagctgctg cctccacttc cctagccagg 120  
 tgctgaagag gatcttcgga gccgctctgg ccccaggcg ctggatgact ggcaccagcg 180  
 ctctcgac ctgtgttgg gtgtgagact tgggctggag tgcccacgtg gctgtggagt 240  
 cagtgtgatt catgattgag gaaacgcgtc ctccatcctc tctctccttg gcactttcca 300  
 cacatgagga gaagaagagc ttctgtttag aagacacgtg ccagagtcga gagggccctt 360  
 gcccacc atg aag gga acc tgt gtt ata gca tgg ctg ttc tca agc ctg 409  
 Met Lys Gly Thr Cys Val Ile Ala Trp Leu Phe Ser Ser Leu →  
 1 5 10  
 ggg ctg tgg aga ctc gcc cac cca gag gcc cag ggt acg act cag tgc 457  
 Gly Leu Trp Arg Leu Ala His Pro Glu Ala Gln Gly Thr Thr Gln Cys  
 15 20 25 30  
 cag aga aca ctc gag gtg aat att gtt tcc ccc agc tcc aag gca aca 505  
 Gln Arg Thr Leu Glu Val Asn Ile Val Ser Pro Ser Ser Lys Ala Thr  
 35 40 45  
 ttc agt cca agt 517  
 Phe Ser Pro Ser  
 50

*move amino acid directly  
 under  
 respective  
 amino acid*

*(please correct this misalignment)  
 in subsequent coding sequences*

*PII* Use of n and/or Xaa has been detected in the Sequence Listing.  
 Review the Sequence Listing to insure a corresponding  
 explanation is presented in the <220> to <223> fields of  
 each sequence using n or Xaa.

VERIFICATION SUMMARY

PATENT APPLICATION: US/09/993,179

DATE: 12/05/2001

TIME: 09:54:31

Input Set : A:\sequence listing.txt

Output Set: N:\CRF3\11212001\I993179.raw

L:12 M:270 C: Current Application Number differs, Replaced Current Application Number  
 L:13 M:271 C: Current Filing Date differs, Replaced Current Filing Date  
 L:41 M:336 W: Invalid Amino Acid Number in Coding Region, SEQ ID:1  
 L:45 M:336 W: Invalid Amino Acid Number in Coding Region, SEQ ID:1  
 L:49 M:336 W: Invalid Amino Acid Number in Coding Region, SEQ ID:1  
 L:59 M:332 E: (32) Invalid/Missing Amino Acid Numbering, SEQ ID:2  
 M:332 Repeated in SeqNo=2  
 L:81 M:336 W: Invalid Amino Acid Number in Coding Region, SEQ ID:3  
 L:85 M:336 W: Invalid Amino Acid Number in Coding Region, SEQ ID:3  
 L:89 M:336 W: Invalid Amino Acid Number in Coding Region, SEQ ID:3  
 L:93 M:336 W: Invalid Amino Acid Number in Coding Region, SEQ ID:3  
 L:97 M:336 W: Invalid Amino Acid Number in Coding Region, SEQ ID:3  
 L:101 M:336 W: Invalid Amino Acid Number in Coding Region, SEQ ID:3  
 L:105 M:336 W: Invalid Amino Acid Number in Coding Region, SEQ ID:3  
 L:109 M:336 W: Invalid Amino Acid Number in Coding Region, SEQ ID:3  
 L:119 M:332 E: (32) Invalid/Missing Amino Acid Numbering, SEQ ID:4  
 M:332 Repeated in SeqNo=4  
 L:142 M:332 E: (32) Invalid/Missing Amino Acid Numbering, SEQ ID:5  
 M:332 Repeated in SeqNo=5  
 L:153 M:332 E: (32) Invalid/Missing Amino Acid Numbering, SEQ ID:6  
 M:332 Repeated in SeqNo=6  
 L:209 M:257 W: Feature value mis-spelled or invalid, <221> Name/Key for SEQ ID#:11  
 L:214 M:341 W: (46) "n" or "Xaa" used, for SEQ ID#:11  
 L:245 M:332 E: (32) Invalid/Missing Amino Acid Numbering, SEQ ID:14  
 M:332 Repeated in SeqNo=14  
 L:256 M:332 E: (32) Invalid/Missing Amino Acid Numbering, SEQ ID:15  
 M:332 Repeated in SeqNo=15